Optimization Problems with Constraints - introduction to theory, numerical Methods and applications

Dr. Abebe Geletu

Ilmenau University of Technology
Department of Simulation and Optimal Processes (SOP)

TECHNISCHE UNIVERSITÄT
ILMENAU

Optimization problems with constraints

A general form of a constrained optimization problem is

$$(NLP) \qquad \min_{x} f(x)$$
$$s.t.$$
$$h_i(x) = 0, i = 1, 2, \ldots, p;$$
$$g_j(x) \leq 0, j = 1, 2, \ldots, m.$$

where $f, g_i, h_j : \mathbb{R}^n \to \mathbb{R}$ are **at least one-time differentiable functions** and $x \in \mathbb{R}^n$.

Feasible set of NLP:

- $\overline{x} \in \mathbb{R}^n$ is **feasible point** of NLP if $h_i(\overline{x}) = 0, i = 1, 2, \ldots, p$ and $g_j(\overline{x}) \leq 0, j = 1, 2, \ldots, m$.
- The set of all feasible points of NLP is the set which we represent by

$$\mathcal{F} := \left\{ x \in \mathbb{R}^n \mid h_i(x) = 0, i = 1, \ldots, p; g_j(x) \leq 0, j = 1, \ldots, m \right\}.$$

known as the **feasible set** of NLP.

Optimization problems with constraints ....

- Some times it is convenient to write the constraints of NLP in a compact form as

$$h(x) = 0, g(x) \leq 0,$$

where

$$h(x) = \begin{pmatrix} h_1(x) \\ h_2(x) \\ \vdots \\ h_p(x) \end{pmatrix} \text{ and } g(x) = \begin{pmatrix} g_1(x) \\ g_2(x) \\ \vdots \\ g_m(x) \end{pmatrix}$$

Example:

$$(NLP1) \qquad \begin{aligned} &\min_x \left\{ \frac{1}{2} x_1^2 + x_1 x_2^2 \right\} \\ &s.t. \\ &\qquad x_1 x_2^2 - 1 = 0, \\ &\qquad -x_1^2 + x_2 \leq 0, \\ &\qquad x_2 \geq 0. \end{aligned}$$

Optimization problems with constraints ....

In the example problem NLP1 above:

- there is only one equality constraint $h_1(x) = x_1 x_2^2 - 1$ and
- two inequality constraints $g_1(x) = -x_1^2 + x_2 \leq 0$ and $g_2(x) = -x_2 \leq 0$.

▶ Observe that the point $x^\top = (1, 1)$ is a feasible point; while the point $(0, 0)$ is not feasible (or **infeasible**); i.e., $x^\top = (0, 0)$ does not belong to the feasible set

$$\mathcal{F} = \left\{ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{R}^2 \mid x_1 x_2^2 - 1 = 0, -x_1^2 + x_2 \leq 0, x_2 \geq 0 \right\}.$$

Optimal Solution

A point $x_* \in \mathbb{R}^n$ is an **optimal solution** of the constrained problem NLP if and only if

(i) if x_* is feasible to NLP; i.e. $x_* \in \mathcal{F}$ and

(ii) $f(x) \geq f(x_*)$ for all $x \in \mathcal{F}$.

Optimization problems with constraints ....

- For the example problem NLP1, the point $x_*^\top = (1, 1)$ is an optimal solution.

▶ In general, it is not easy to determine an optimal solution of a constrained optimization problem like NLP.

Question:

(**Q1**) How do we verify that a give point $\overline{x} \in \mathbb{R}^n$ is an optimal solution of NLP? (This about optimality criteria.)
(**Q2**) What methods are available to find an optimal solution of NLP? (This about computation of optimal solutions.)

Definition (Active Constraints)

- Let $\overline{x}$ be a feasible point of NLP. An inequality constraint $g_i(x) \le 0$ of NLP is said to be **active** at $\overline{x}$ if

$$g_i(\overline{x}) = 0.$$

- The set $\mathcal{A}(\overline{x}) = \{i \in \{1, 2, \ldots, m\} \mid g_i(\overline{x}) = 0\}$ denotes **index set of active constraints** at $\overline{x}$.

Optimization with constraints ... Optimality criteria

- For the example problem NLP1, the constraint $g_1 = -x_1^2 + x_2$ is active at the point $x_*^\top = (1,1)$; while $g_2 = -x_2$ is active at x_* . Hence, $\mathcal{A} = \{1\}$

descent direction

A vector d is a **descent direction** for the objective function f at the point $\overline{x}$ if

$$f(\overline{x} + d) \leq f(\overline{x})$$

Moving from point $\overline{x}$ in the direction of d decreases the function f.

- **Any vector d that satisfies $d^\top f(\overline{x}) \leq 0$ is a descent direction.**

To verify this, recall the 1st order Taylor approximation of f at the point $\overline{x}$.

$$f(\overline{x} + d) \approx f(\overline{x}) + d^\top \nabla f(\overline{x}).$$

It follows that

$$f(\overline{x} + d) - f(\overline{x}) = d^\top \nabla f(\overline{x}) \leq 0 \Rightarrow f(\overline{x} + d) - f(\overline{x}) \leq 0.$$

Hence, $f(\overline{x} + d) \leq f(\overline{x})$. Therefore, d is a descent direction.

▶ **NB:** If d is a decent direction, then $\widetilde{d} = \alpha d, \alpha > 0$, is also a descent direction.

Optimization with constraints ... Optimality criteria

feasible direction

Let $\overline{x}$ a feasible point (i.e. $\overline{x} \in \mathcal{F}$) and d is a vector in $\mathbb{R}^n$. If
(i) $h_i(\overline{x}+d)=0, i=1,\ldots,p$ and
(ii) $g_j(\overline{x}+d)\leq 0, j=1,\ldots,m$,
then d is said to be to be a **feasible direction** for the NLP.

- **Let $\overline{x}$ a feasible point. If a vector d satisfies**

$$d^\top \nabla h_i(\overline{x})=0, i=1,\ldots,m \textbf{ and } d^\top \nabla g_j(\overline{x})<0, j\in\mathcal{A}(\overline{x}),$$

then $\widetilde{d}=\alpha d$ is feasible direction at $\overline{x}$, for some $\alpha>0$.

To verify this,use the 1st order Taylor approximations at the point $\overline{x}$.

$$i=1,\ldots,p: \quad h_i(\overline{x}+\alpha d) \approx \underbrace{h_i(\overline{x})}_{=0}+\alpha \underbrace{d^\top \nabla h_i(\overline{x})}_{=0}; \Rightarrow \underline{h_i(\overline{x}+\widetilde{d})=0},$$

$$j\in\mathcal{A}(\overline{x}): \quad g_j(\overline{x}+\alpha d) \approx \underbrace{g_j(\overline{x})}_{\leq 0}+\alpha \underbrace{d^\top \nabla g_j(\overline{x})}_{<0} \Rightarrow \underline{g_j(\overline{x}+\widetilde{d})\leq 0};$$

$$\underbrace{j\notin\mathcal{A}(\overline{x})}_{\text{non active constraints}}: \quad g_j(\overline{x}+\alpha d) \approx \underbrace{g_j(\overline{x})}_{<0}+\alpha d^\top \nabla g_j(\overline{x}) \Rightarrow \underline{g_j(\overline{x}+\widetilde{d})\leq 0}, \text{ for } 0<\alpha\leq -\frac{g_j(\overline{x})}{d^\top \nabla g_j(\overline{x})} \text{ if } d^\top \nabla g_j(\overline{x})>0.$$

Optimization with constraints ... Optimality criteria

Optimality condition

If x_* is an optimal solution of NLP, then there is no vector $d \in \mathbb{R}^n$ which is both a descent direction for f and feasible direction at x_*.

That is, if x_* an optimal solution of NLP, then system of inequalities

$$d^\top \nabla f(x_*) < 0,\ d^\top \nabla h_i(x_*) = 0,\ i = 1, \dots, p;\ d^\top \nabla g_j(x_*) < 0,\ j \in \mathcal{A}(x_*)$$

equivalently

$$[-\nabla f(x_*)]^\top d > 0,\ [\nabla h_i(x_*)]^\top d = 0,\ i = 1, \dots, p;\ \left[\nabla g_j(x_*)\right]^\top d < 0,\ j \in \mathcal{A}(x_*) \quad (1)$$

has no solution d.

Farkas' Theorem

Given any set of vectors $c, a_i, b_j \in \mathbb{R}^n$, $i = 1, \dots, m; j = 1, \dots, l$. Then **exactly only one of the following two systems has a solution**

System I: $c^\top d > 0,\ a_i^\top d = 0,\ i = 1, \dots, p;\ b_j^\top d < 0,\ j = 1, \dots, \widetilde{m}$

System II: $c = \sum_{i=1}^{p} \lambda_i a_i + \sum_{j=1}^{l} \mu_j b_j,\ \mu > 0.$

Now if we let $c = -\nabla f(x_*)$, $a_i = \nabla h_i(x_*)$, $i = 1, \dots, p$ and $b_j = \nabla g_j(x_*)$, $j \in \mathcal{A}(x_*)$, then, since x_* an optimal point, then only system II has a solution.

Optimization with constraints ... Optimality criteria

- Hence, if x_* is an optimal solution of NLP, then there exist vectors $\lambda \in \mathbb{R}^m$, $\lambda^\top = (\lambda_1, \lambda_2, \ldots, \lambda_m)$ and $\mu \in \mathbb{R}^l$, $\mu^\top = (\mu_1, \mu_2, \ldots, \mu_l) > 0$ such that

$$-\nabla f(x_*) = \sum_{i=1}^{m} \lambda_i \nabla h_i(x_*) + \sum_{j=1}^{l} \mu_j \nabla g_j(x_*)$$

where $l = \#\mathcal{A}(x_*)$. If we let $\mu_j = 0$ for for $j \in \{1, \ldots, m\} \setminus \mathcal{A}(x_*)$, then we can write $-\nabla f(x_*) = \sum_{i=1}^{m} \lambda_i \nabla h_i(x_*) + \sum_{j=1}^{m} \mu_j \nabla g_j(x_*)$.

The Karush-Kuhn-Tucker (KKT) optimality condition

If x_* is a minimum point of NLP, then there is $\lambda \in \mathbb{R}^p$ and $\mu \in \mathbb{R}^m, \mu \geq 0$ such that the following hold true:

$$\begin{aligned}
\nabla f(x_*) + \sum_{i=1}^{m} \lambda_i \nabla h_i(x_*) + \sum_{j=1}^{m} \mu_j \nabla g_j(x_*) &= 0 && \text{(Optimality)} \\
h(x_*) &= 0 && \text{(feasibility)} \\
g(x_*) &\leq 0 && \\
\mu &\geq 0 && \text{(Nonnegativity)} \\
\mu_j g_j(x_*) &= 0, j = 1, \ldots, m. && \text{(Complementarity)}
\end{aligned}$$

Optimization with constraints ... Optimality criteria

- The function

$$\mathcal{L}(x,\lambda,\mu) = f(x) + \sum_{i=1}^{m} \lambda_i h_i(x) + \sum_{j=1}^{m} \mu_j g_j(x)$$

called the **Lagrange function** associated to the NLP.

Example: Solve the following optimization problem:

$$
\begin{aligned}
(NLP2) \qquad & \min_x \left\{ f(x) = x_1^2 - x_2^2 \right\} && (2)\\
& s.t. && (3)\\
& \qquad x_1 + 2x_2 + 1 = 0 && (4)\\
& \qquad x_1 - x_2 \leq 3. && (5)
\end{aligned}
$$

Solution:

Lagrange function

$\mathcal{L}(x,\lambda,\mu) = (x_1^2 - x_2^2) + \lambda(x_1 + 2x_2 + 1) + \mu(x_1 - x_2 - 3).$

Optimization with constraints ... Optimality criteria

Optimality condition:

$$\frac{\partial \mathcal{L}}{\partial x_1} = 0 \Rightarrow 2x_1 + \lambda + \mu = 0 \Rightarrow x_1 = -\frac{1}{2}(\lambda + \mu) \tag{6}$$

$$\frac{\partial \mathcal{L}}{\partial x_2} = 0 \Rightarrow -2x_2 + 2\lambda - \mu = 0 \Rightarrow x_2 = \frac{1}{2}(2\lambda - \mu) \tag{7}$$

Feasibility

$$\begin{aligned} h(x) = 0 \Rightarrow x_1 + 2x_2 + 1 = 0 \quad &\Rightarrow \quad -\frac{1}{2}(\lambda + \mu) + (2\lambda - \mu) + 1 = 0 \\ &\Rightarrow \quad \lambda = \mu - \frac{2}{3} \end{aligned}$$

Complementarity

$$\begin{aligned} \mu g(x) = 0 \quad &\Rightarrow \quad \mu(x_1 - x_2 - 3) = 0 \Rightarrow \mu\left(-\frac{1}{2}(\lambda + \mu) - \frac{1}{2}(2\lambda - \mu) - 3\right) = 0 \\ &\Rightarrow \quad \mu\left(-\frac{1}{2}\left[(\mu - \frac{2}{3}) + \mu\right] - \left[(\mu - \frac{2}{3}) - \mu\right] - 3\right) = 0 \\ &\Rightarrow \quad \mu\left[-\mu - 2\right] = 0 \Rightarrow \mu = 0 \text{ or } \mu = -2. \end{aligned}$$

$\mu = -2 < 0$ is not allowed. Hence, $\mu_* = 0$ is the only possibility. As a result $\lambda_* = \mu_* - \frac{2}{3} = -\frac{2}{3}$.

Optimization with constraints ... Optimality criteri

Now using $\mu_* = 0$ and $\lambda_* = -1$

$$x_1^* = -\frac{1}{2}(\lambda_* + \mu_*) = -\frac{1}{2}(-\frac{2}{3} + 0) = \frac{1}{3} \tag{8}$$

$$x_2^* = \frac{1}{2}(2\lambda_* - \mu_*) = \frac{1}{2}(2 \times (-\frac{2}{3}) - 0) = -\frac{2}{3}. \tag{9}$$

The point $x^{*\top} = \left(\frac{1}{3}, -\frac{2}{3}\right)$ is a candidate for optimality.

Observe that: the inequality constraint $g(x) = x_1 - x_2 - 3 \leq 0$ is not active at x^*.

- In general, the KKT conditions above are only necessary optimality conditions.

Optimization with constraints ... Optimality criteria

Sufficient optimality conditions

Suppose the functions $f, h_i, i =, \ldots, p; g_j, j = 1, \ldots, m$ are twice differentiable. Let x^* be a feasible point of NLP. If there are Lagrange multipliers λ^* and $\mu^* \geq 0$ such that:
(i) the KKT conditions are satisfied at (x^*, μ^*, λ^*); and
(ii) and the hessian of the Lagrangian

$$\nabla_x L(x^*, \lambda^*, \mu^*) = \nabla^2 f(x^*) + \sum_{i=1}^{p} \lambda_i^* \nabla^2 h_j(x^*) + \sum_{j=1}^{m} \mu_j^* \nabla^2 g_j(x^*)$$

is positive definite (i.e. $d^\top \nabla_x L(x^*, \lambda^*, \mu^*) d > 0$) for d from the subspace $\mathcal{S} = \left\{ d \in \mathbb{R}^n \mid d^\top \nabla h_i(x^*) = 0, i = 1, \ldots, p; d^\top \nabla g_j(x^*) = 0, \mu_j > 0, j \in \mathcal{A}(x^*) \right\}$, then x^* is an optimal solution of NLP.

For the problem NLP2, since $g(x^*) < 0$ we have

- $\mathcal{S} = \{d \in \mathbb{R}^2 \mid d^\top \nabla h(x_*) = 0\} = \{d \in \mathbb{R}^2 \mid (d_1, d_2) \begin{pmatrix} 1 \\ 2 \end{pmatrix} = 0\} = \{d \in \mathbb{R}^2 \mid d_1 = -2d_2\}$.
- $\nabla_x L(x^*, \lambda^*, \mu^*) = \begin{pmatrix} 2 & 0 \\ 0 & -2 \end{pmatrix}$. Let $d^\top = (d_1 . d_2)^\top \in \mathcal{S}$, $d \neq 0$ (note that if $d_1 \neq 0$, then $d_2 \neq 0$ and vice-versa.)

$$(d_1, d_2) \begin{pmatrix} 2 & 0 \\ 0 & -2 \end{pmatrix} \begin{pmatrix} d_1 \\ d_2 \end{pmatrix} = (-2d_2, d_2) \begin{pmatrix} 2 & 0 \\ 0 & -2 \end{pmatrix} \begin{pmatrix} -2d_2 \\ d_2 \end{pmatrix} = 2d_2 > 0.$$

Therefore, $x^{*\top} = \left(\frac{1}{3}, -\frac{2}{3} \right)$ is an optimal solution of NLP2.

Numerical methods for constrained optimization problems

▶ Unless in some trivial cases, solution of a constrained optimization problem by hand is not a trivial task.

• As a result, currently we find several numerical methods to determine an approximate solution(s) for a constrained optimization problem. Some well known gradient-based methods are

(I) The sequential quadratic programming (SQP) method
(II) The Interior Point Method
(III) Penalty function methods
(IV) The Augmented Lagrangian method, etc

▶ SQP based algorithms are highly favored for the solution of constrained nonlinear optimization problems.

▶ In the SQP method: a sequence of iterates $x_{k+1} = x_k + \alpha_k d_k$ are generated by solving **quadratic optimization (programming)** problems to determine the d_k's.

Quadratic optimization (programming) problems

• Quadratic optimization problems appear in several engineering applications.
Eg: Least square approximations (regression), Model predictive control, in SQP method for NLP, etc.

(I) Quadratic programming problems with equality constraints

$$(QP) \qquad \min_{x} \left\{ f(x) = x^\top Q x + q^\top x \right\} \tag{10}$$

$$s.t. \tag{11}$$

$$Ax = b, \tag{12}$$

where:
- $Q \in \mathbb{R}^{n \times n}$ a symmetric (i.e. $Q^\top = Q$) positive definite matrix.
- $x \in \mathbb{R}^n$ the (unknown) optimization variable; $q \in R^n$ a given vector.
- $A \in \mathbb{R}^{m \times n}$ and $b \in \mathbb{R}^m$ are given; i.e. there are m equality constraint.
- $rank(A) = m$.

Quadratic optimization ....

The Lagrange function for (QP):

$$\mathcal{L}(x,\lambda) = x^\top Qx + q^\top x + \lambda^\top (Ax - b)\,.$$

Karush-Kuhn-Tucker Conditions:

$$\frac{\partial \mathcal{L}}{\partial x} = 0 \Rightarrow Qx + A^\top \lambda = -q \tag{13}$$

$$\frac{\partial \mathcal{L}}{\partial \lambda} = 0 \Rightarrow Ax = b. \tag{14}$$

This implies

$$\begin{bmatrix} Q & A^\top \\ A & 0 \end{bmatrix} \begin{bmatrix} x \\ \lambda \end{bmatrix} = \begin{bmatrix} -q \\ c \end{bmatrix}$$

- If Q is positive definite and $rank(A) = m$, then $\begin{bmatrix} Q & A^\top \\ A & 0 \end{bmatrix}$ is invertible and

$$\begin{bmatrix} x \\ \lambda \end{bmatrix} = \begin{bmatrix} Q & A^\top \\ A & 0 \end{bmatrix}^{-1} \begin{bmatrix} -q \\ c \end{bmatrix}$$

$$\Rightarrow \qquad x^* = -Q^{-1}q + Q^{-1}A^\top \left(AQ^{-1}A^\top\right)^{-1} \left(b + AQ^{-1}q\right)$$

$$\lambda^* = -\left(AQ^{-1}A^\top\right)^{-1} \left(b + AQ^{-1}q\right).$$

Quadratic optimization (programming) problems

Example:

A material flow from two reservoirs R1 and R2 flows into a third reservoir R3. The rate of flow from R1 is measured to be $F_1^M = 6.5m^3/h$ and from R2 is $F_2^M = 14.7m^3/h$. As a result of these two inflows, the volume in R3 is measured to change at a rate $F_3^M = 19.8m^3/h$.

Flow rate measurement devices for each reservoirs are known to incur measurement errors each with standard-deviation: $\sigma_1 = 0.1m^3/h$, $\sigma_2 = 0.2m^3/h$ and $\sigma_3 = 0.3m^3/h$.

State an optimization problem to determine the actual flow rates.

Solution: Let F_1, F_2 and F_3 be the true flow rates.

- Note that $F_i = F_i^M \pm \eta_i \sigma_i, i = 1, 2, 3;$

$\Rightarrow$ measurement $\pm\eta_i\sigma_i = F_i - F_i^M$ represents measurement error.

$\Rightarrow$ To minimize all these measurement error

$$\min_{(F_1,F_2,F_3)} \left\{ \left(\frac{F_1 - F_1^M}{\sigma_1} \right) + \left(\frac{F_2 - F_2^M}{\sigma_2 2} \right) + \left(\frac{F_3 - F_3^M}{\sigma_3} \right) \right\}$$
$$s.t.$$
$$F_1 + F_2 = F_3.$$

Quadratic optimization (programming) problems

(II) Quadratic programming problems with equality and inequality constraints

$$(QP) \qquad \min_{x} \left\{ f(x) = \mathbf{x}^\top Q\mathbf{x} + \mathbf{q}^\top x \right\} \tag{15}$$

$$s.t. \tag{16}$$

$$\mathbf{a}_i^\top x = b_i, i = 1, 2, \ldots, p \tag{17}$$

$$\mathbf{a}_i x \geq b_i, i = p+1, \ldots, m. \tag{18}$$

where: Q is symmetric positive definite.
If we set

$A = [\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_p], \mathbf{a}^\top = (b_1, b_2, \ldots, b_p), B = [\mathbf{a}_{p+1}, \mathbf{a}_{p+2}, \ldots, \mathbf{a}_m]$ and $\mathbf{b}^\top = (b_{p+1}, b_{p+2}, \ldots, b_m)$

Then (QP) can be compactly written as:

$$(QP) \qquad \min_{x} \left\{ f(x) = \mathbf{x}^\top Q\mathbf{x} + \mathbf{q}^\top x \right\} \tag{19}$$

$$s.t. \tag{20}$$

$$A\mathbf{x} = \mathbf{a} \tag{21}$$

$$B\mathbf{x} \geq \mathbf{b}. \tag{22}$$

- In general, inequality constrained (QP)s may not be easy to be solved analytically.

Active Set Method

Strategy:

- Start from an arbitrary point x^0
- Find the next iterate by setting $x^{k+1} = x^k + \alpha_k d^k$,

where α_k is a step-length and d^k is search direction.

Question

- How to determine the search direction d^k?
- How to determine the step-length α_k?

(A) **Determination of the search direction:**

- At the current iterate x^k determine the index set of active constraints

$$\mathcal{A}^k = \{i \mid a_i x^k - b_i = 0, i = p+1, \ldots, m\} \cup \{1, 2, \ldots, p\}.$$

Active Set Method

Strategy:

- Start from an arbitrary point x^0
- Find the next iterate by setting $x^{k+1} = x^k + \alpha_k d^k$,

where α_k is a step-length and d^k is search direction.

Question

- How to determine the search direction d^k?
- How to determine the step-length α_k?

(A) **Determination of the search direction:**

- At the current iterate x^k determine the index set of active constraints

$$\mathcal{A}^k = \{i \mid a_i x^k - b_i = 0, i = p+1, \ldots, m\} \cup \{1, 2, \ldots, p\}.$$

Quadratic optimization ... Active set Method

- Solve the direction finding problem

$$\min_d \left\{ \frac{1}{2}(x^k + d)^\top Q(x^k + d) + q^\top (x^k + d) \right\}$$
$$s.t.$$
$$a_i^\top (x^k + d) = b_i, i \in \mathcal{A}^k.$$

Expand

- $\frac{1}{2}(x^k + d)^\top Q(x^k + d) + q^\top (x^k + d) = \frac{1}{2} d^\top Qd + \underline{\frac{1}{2} d^\top Qx^k + \frac{1}{2}(x^k)^\top Qd} + \frac{1}{2}(x^k)^\top Qx^k + q^\top d + q^\top x^k$
- $a_i \top \left(x^k + d\right) = b_i \Rightarrow a_i \top d = b_i - a_i^\top x^k = 0.$
- Simplify these expressions and drop constants to obtain:

$$\min_d \left\{ \frac{1}{2} d^\top Q + d + \left[Qx^k + q \right] d \right\}$$
$$s.t.$$
$$a_i^\top d = 0, i \in \mathcal{A}^k.$$

Set

$$A = \begin{bmatrix} a_1^\top \\ a_2^\top \\ \vdots \end{bmatrix}, i \in \mathcal{A}^k.$$

- Solve this problem using the problem using the KKT conditions to obtain d_k.

Quadratic optimization ... Active set Method

(A) **Determination of the step length:**
Once you obtained d_k compute α_k as

$$\alpha_k = \min \left[1, \frac{b_i - a_i^\top x^k}{a_i^\top d^k}, i \notin \mathcal{A}^k \right]$$

Quadratic optimization ... Active set Method

Example: Application Problem

A parallel flow heat exchanger of a given length to be designed. The conducting tubes, all of the same diameter, are enclosed in an outer shell of diameter $D = 1m$ (see Fugure). The diameter of each of the conducting tubes should not exceed $60mm$. Determine the number of tubes and the diameter for the largest surface area.